# Answer Sheet for Homework 1

2.6. The table shows that        

(a) 

(b) 

(c) Calculate the conditional probabilities first:



The conditional expectations are



Use the solution to part (b),

Unemployment rate for college graduates  1  *E*(*Y*|*X*  1)  1  0.974  0.026

Unemployment rate for non-college graduates  1  *E*(*Y*|*X*  0)  1  0.944  0.056

(d) The probability that a randomly selected worker who is reported being unemployed is a college graduate is



The probability that this worker is a non-college graduate is



(e) Educational achievement and employment status are not independent because they do not satisfy that, for all values of *x* and *y*,



For example, from part (e)  while from the table Pr(*X*  0)  0.659.

2.10. Using the fact that if  then  and Appendix Table 1, we have

(a) 

(b) 

(c) 

(d) 

2.13. (a) 

(b) *Y* and *W* are symmetric around 0, thus skewness is equal to 0; because their mean is zero, this means that the third moment is zero.

(c) The kurtosis of the normal is 3, so  ; solving yields  a similar calculation yields the results for *W*.

(d) First, condition on  so that 



Similarly,



From the law of iterated expectations



(e)  thus from part (d). Thus skewness  0. Similarly,  and  Thus,

2.23. *X* and *Z* are two independently distributed standard normal random variables, so

(a) Because of the independence between  and   and  Thus 

(b)  and 

(c)  Using the fact that the odd moments of a standard normal random variable are all zero, we have  Using the independence between  and  we have  Thus 

(d) 

2.26. (a) corr(*Yi*,*Yj*)  , where the first equality uses the definition of correlation, the second uses the fact that *Yi* and *Yj* have the same variance (and standard deviation), the third equality uses the definition of standard deviation, and the fourth uses the correlation given in the problem. Solving for cov(*Yi*, *Yj*) from the last equality gives the desired result.

(b) , so that *E*()  

(c) , so that 



where the fourth line uses  for any variable *a*.

(d) When *n* is large  and , and the result follows from (c).

3.3. Denote each voter’s preference by   if the voter prefers the incumbent and  if the voter prefers the challenger.  is a Bernoulli random variable with probability Pr and Pr From the solution to Exercise 3.2,  has mean  and variance 

(a) 

(b) The estimated variance of  is . The standard error is

(c) The computed *t*-statistic is  Because of the large sample size  we can use Equation (3.14) in the text to get the  
*p*-value for the test  vs.  

(d) Using Equation (3.17) in the text, the *p*-value for the test  vs.  is 

(e) Part (c) is a two-sided test and the *p*-value is the area in the tails of the standard normal distribution outside ± (calculated *t*-statistic). Part (d) is a one-sided test and the *p*-value is the area under the standard normal distribution to the right of the calculated *t*-statistic.

(f) For the test  vs.  we cannot reject the null hypothesis at the 5% significance level. The *p*-value 0.066 is larger than 0.05. Equivalently the calculated *t*-statistic  is less than the critical value 1.64 for a one-sided test with a 5% significance level. The test suggests that the survey did not contain statistically significant evidence that the incumbent was ahead of the challenger at the time of the survey.

3.9. Denote the life of a light bulb from the new process by  The mean of  is  and the standard deviation of  is  hours.  is the sample mean with a sample size  The standard deviation of the sampling distribution of  is  hours. The hypothesis test is  vs.  The manager will accept the alternative hypothesis if  hours.

(a) The size of a test is the probability of erroneously rejecting a null hypothesis when it is valid.

The size of the manager’s test is



where  means the probability that the sample mean is greater than 2100 hours when the new process has a mean of  hours.

(b) The power of a test is the probability of correctly rejecting a null hypothesis when it is invalid. We calculate first the probability of the manager erroneously accepting the null hypothesis when it is invalid:



The power of the manager’s testing is 

(c) For a test with 5%, the rejection region for the null hypothesis contains those values of the   
*t*-statistic exceeding 1.645.



The manager should believe the inventor’s claim if the sample mean life of the new product is greater than 2032.9 hours if she wants the size of the test to be 5%.

3.11. Assume that  is an even number. Then  is constructed by applying a weight of 1/2 to the *n*/2 “odd” observations and a weight of 3/2 to the remaining *n*/2 observations.

